WO0067580

Publication Title:

METHOD FOR IDENTIFYING COMPOUNDS THAT MODULATE EUKARYOTIC SPLICING

Abstract:

Abstract of WO0067580

The invention relates to a method of screening for compounds that modulate eukaryotic splicing reactions, using assays based on such splicing reactions. Also provided are methods of using in <i>in vivo</i> or <i>in vitro</i> splicing reactions to screen for compounds that modulate splicing reactions. Data supplied from the esp@cenet database - Worldwide

---

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
A01N 61/00, A61K 31/00, C12P 19/34, C12Q 1/68

A1

(11) International Publication Number: **WO 00/67580**

(43) International Publication Date: 16 November 2000 (16.11.00)

(21) International Application Number: PCT/US00/12249

(22) International Filing Date: 4 May 2000 (04.05.00)

(30) Priority Data:
60/133,001 7 May 1999 (07.05.99) US

(71) Applicants *(for all designated States except US)*: SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US). SMITHKLINE BEECHAM PLC [GB/GB]; New Horizons Court, Great West Road, Brentford, Middlesex TW8 9EP (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: GONTAREK, Richard, R. [US/US]; 511 Hagey Place, Trappe, PA 19426 (US).

(74) Agents: GIMMI, Edward, R. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406–0939 (US).

(81) Designated States: JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: METHOD FOR IDENTIFYING COMPOUNDS THAT MODULATE EUKARYOTIC SPLICING

(57) Abstract

The invention relates to a method of screening for compounds that modulate eukaryotic splicing reactions, using assays based on such splicing reactions. Also provided are methods of using in *in vivo* or *in vitro* splicing reactions to screen for compounds that modulate splicing reactions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHOD FOR IDENTIFYING COMPOUNDS THAT MODULATE EUKARYOTIC SPLICING

FIELD OF THE INVENTION

This invention relates to a method of screening for compounds that modulate gene expression, particularly those that modulate a splicing reaction, using assays based on eukaryotic splicing reactions. Also provided are methods of using in *in vivo* or *in vitro* splicing reactions to screen for compounds that modulate the splicing reaction. Further provided are methods of treating an individual with such compounds to treat diseases, particularly infections caused by or related to eukaryotic pathogens.

BACKGROUND OF THE INVENTION

It is well known that eukaryotic genes are often interrupted by introns that must be removed by splicing in the cell nucleus before the mature mRNA can be translated in the cytoplasm. This pre-mRNA splicing requires a number of protein, small nuclear ribonucleoprotein (herein "snRNP(s)") factors and pre-mRNA sequence signals (herein "splicing signals," "splicing sequences" or "splicing polynucleotides") to form the spliceosome, which is the large macromolecular complex that catalyzes an efficient and accurate removal of introns. Although the overall mechanism of splicing is conserved between yeast and man, there exists divergence in the splicing signals and the soluble factors required. More than 37 precursor RNA processing (herein "PRP") genes have been identified in yeast whose products affect splicing. Many of these are essential factors. It is believed that fungal splicing signals are conserved in certain pathogenic fungi (*e.g.*, *Candida* and *Aspergillus*). Moreover, soluble *in vitro* splicing systems are well characterized for *S. cerevisiae*. In addition, there are secondary assays available for dissecting the splicing reaction (such as native gels for looking at splicing complex formation, among others known in the art).

The study of pre-mRNA splicing in eukaryotes, in particular in the yeasts, and especially in *Saccharomyces cerevisiae* has seen remarkable advances in recent years largely due to the development of the soluble, *in vitro* splicing systems. However, as yet there is an unmet medical need to understand splicing in pathogens, particularly to understand it in a manner that facilitates the development of new drugs. This invention provides both steps towards understanding pre-mRNA splicing in pathogenic eukaryotic

taxa, in particular in fungal species, as well as providing assays and compounds useful in the discovery of drugs against eukaryotic pathogens.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

SUMMARY OF THE INVENTION

The invention provides a method of screening for a compound modulating a pathogen or non-pathogen splicing reaction comprising the steps of contacting a splicing reaction mix with a candidate compound and identifying a compound modulating the splicing reaction.

The invention also provides a method wherein the modulation is inhibition.

In a still a further embodiment of the invention there is provided a method wherein the modulating activity is activation.

In yet a further embodiment of the invention there is provided a method wherein the pathogen or non-pathogen is selected from the group consisting of a fungus, protozoan, nematode and arthropod.

In a further embodiment of the invention there is provided a method of wherein the splicing reaction comprises at least one component of a pathogen or non-pathogen spliceosome.

In yet a further embodiment of the invention there is provided a method of screening for a compound modulating a pathogen or non-pathogen splicing reaction comprising the steps of contacting a splicing reaction mix with a candidate compound and detecting a compound modulating the splicing reaction.

Another embodiment of the invention provides compound identified using the method of the invention or selected using a method of the invention.

In a further embodiment of the invention there is provided a method of inhibiting a splicing reaction comprising the step of contacting the reaction machinery with a compound of the invention.

In still a further embodiment of the invention there is provided a method of inhibiting a splicing reaction of the invention wherein the reaction is in a pathogen or non-pathogen.

In a further embodiment of the invention there is provided a method of treating an individual infected with a pathogen comprising the step of contacting the individual with the a compound of the invention

In a further embodiment of the invention there is provided a method of treating an individual infected with a pathogen comprising the step of contacting the individual with the a compound that modulates the activity of a splicing reaction in a pathogen or non-pathogen.

In a further embodiment of the invention there is provided an antagonist or agonist of a pathogen or non-pathogen splicing reaction.

In a further embodiment of the invention there is provided a method of the invention wherein the fungus is selected form the group consisting of *S. cerevisiae* and *C. albicans*.

In a further embodiment of the invention there is provided a method of screening for a compound modulating a the maturation of pathogen or non-pathogen pre-mRNA or hn-RNA to mRNA comprising the steps of contacting a composition comprising a pre-mRNA or hn-RNA with a candidate compound and identifying a compound modulating the maturation of pathogen or non-pathogen pre-mRNA or hn-RNA to mRNA.

In a further embodiment of the invention there is provided a compound identified using a method of the invention.

In a further embodiment of the invention there is provided a method of modulating the maturation of pathogen or non-pathogen pre-mRNA or hn-RNA to mRNA with a compound of the invention.

In a further embodiment of the invention there is provide a method of treating an individual infected with a pathogen comprising the step of contacting the individual with the a compound identified using a method of the invention.

In a further embodiment of the invention there is provided a method of treating an individual infected with a pathogen comprising the step of contacting the individual with a compound that modulates the maturation of pre-mRNA or hn-RNA to mRNA.

In a further embodiment of the invention there is provided a method of treating an individual infected with a pathogen comprising the step of contacting the individual with the compound of the invention.

GLOSSARY

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Disease(s)" means any infection or illness caused by, suspected of being caused by, or associated with a pathogen or pathogens (as "pathogen(s)" is defined herein).

"Host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM *J. Applied Math., 48:* 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research 12(1):* 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S.F. et al., *J. Molec. Biol. 215:* 403-410 (1990). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, MD 20894; Altschul, S., *et al., J. Mol. Biol. 215*: 403-410 (1990). The well-known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:

1) Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)

Gap Penalty: 12

Gap Length Penalty: 4

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:

1) Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default parameters for nucleic acid comparisons.

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to the reference sequence of SEQ ID NO:1 OR 2, wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO: 1 or may include up to a certain integer number of nucleotide alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5''terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 OR 2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1 OR 2, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ ID NO:1 OR 2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and • is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n. Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:1 OR 2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1 OR 2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one nucleic acid deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at

the 5''terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of amino acids in SEQ ID NO:1 OR 2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:1 OR 2, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of amino acid alterations, x_n is the total number of amino acids in SEQ ID NO:1 OR 2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., • is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n.

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50,60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:1 OR 2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO: 2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:1 OR 2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:1 OR 2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:1 OR 2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and • is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a.

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1 OR 2, that is it may be 100% identical, or it may

include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:1 OR 2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:1 OR 2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:1 OR 2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and • is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

The methods of the invention may be used with a wide variety of microbes to screen for compounds that are antimicrobial compounds, such as compounds that are microbistatic or microbicidal. Herein "microbe(s)" and "microbial" means any microscopic and/or unicellular fungus, any bacteria, and any protozoan.

"Pathogen(s)" means any organism that is known to, or capable or suspected of causing infection in a vertebrate, which organisms include but not limited to a nematode, arthropod, or unicellular or filamentous eukaryote, including but not limited to, a protozoan, a fungus, a member of the genus, *Aspergillus, Coccidiodes, Histoplasma, Cryptococcus, Paracoccidioides,* or *Candida*, including, but not limited to a member of the species *Kluveromyces lactis*, or *Candida albicans*.

"Non-pathogen(s)" means any organism that is not capable of causing, or is not known to or suspected of causing infection in a vertebrate, which organisms include but not limited to a nematode, arthropod, or unicellular or filamentous eukaryote, including but not

limited to, a protozoan, a fungus, a member of the genus, *Saccharomyces*, including, but not limited to a member of the species *Saccharomyces ceriviseae*.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotide(s)" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded regions, or a mixture of single- and double-stranded regions. In addition, "polynucleotide" as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term "polynucleotide(s)" also includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotide(s)" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term "polynucleotide(s)" as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including, for example, simple and complex cells. "Polynucleotide(s)" also embraces short polynucleotides often referred to as oligonucleotide(s).

"Polypeptide(s)" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" refers to both short chains, commonly referred to as peptides, oligopeptides and oligomers and to longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification

may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. Modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins, such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993) and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990) and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

"pre-mRNA(s)" means any natural, artificial or hybrid natural-artificial RNA molecule that mediates a splicing reaction *in vitro*, *in vivo* or in a splicing reaction mix.

"Splicing reaction(s)" means any or all steps in the maturation of a eukaryotic pre-mRNA, including but not limited to lariat formation, branchpoint formation, spliceosome formation (or any aggregation of splicing factors or proteins on an mRNA, pre-mRNA or polynucleotide splicing signal), PRP binding or expression, snRNP-RNA interactions, binding of any splicing factor or protein bound to any RNA sequence that is, or is derived from, a pre-mRNA any sequence.

"Splicing reaction mix(es)" means a composition that allows a pre-mRNA to undergo a splicing reaction.

"Splicing reaction machinery" means factors and proteins that allows a pre-mRNA to undergo a splicing reaction, including, but not limited to, an intron, a lariat structure, a branchpoint sequence, a spliceosome or any component thereof, any aggregation of splicing factors or proteins capable of binding an mRNA, pre-mRNA or polynucleotide splicing signal, a PRP, an snRNP-RNA or any component thereof, any splicing factor or protein capable of binding to an RNA sequence that is, or is derived from, a pre-mRNA any sequence.

"Variant(s)" as the term is used herein, is a polynucleotide, such as a pre-mRNA, or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide.. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a diagram of the two pre-mRNA splicing substrates is shown (exons are represented as boxes, and introns as lines).

Figure 2 shows a *C. albicans* beta-Tub pre-mRNA accurately spliced in *in vitro*, 15 minute and 45 minute splicing reactions performed and analyzed via polyacrylamide gel electrophoresis.

Figure 3 shows that both Act1 and beta-Tub pre-mRNAs form splicing complexes in cell extracts from both *S. cerevisiae* and *C. albicans*. A timecourse of *in vitro* splicing was performed in *C. albicans* whole cell extracts as described in the Examples.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides novel methods and compositions useful for modulating splicing reactions, particularly those reactions in pathogens.

Also provided herein are methods and compounds useful in assays to screen for compounds that can modulate splicing reactions in pathogens and non-pathogens, particularly unicellular fungi.

Further embodiments of the invention provide high throughput screening methods (herein "HTS") to screen for agonists or antagonists of splicing reactions.

The reactions and signals involved in pre-mRNA splicing provide a number of useful targets for the discovery of compound that can modulate the splicing reaction and can be used anti-pathogen drugs, particularly antifungals drugs. Since there exists genetic divergence in splicing signals and the soluble factors required, the present invention provides methods by which this global process can be targeted specifically to find compounds that can modulate splicing activity.

The invention also provides a method of screening for a compound modulating a pathogen or non-pathogen splicing reaction comprising the steps of contacting a splicing reaction mix with a candidate compound and identifying a compound modulating the splicing reaction. Preferred reaction mixes comprise factors and proteins required for a pre-mRNA to undergo a splicing reaction. Such proteins and factors are known in the art. Most preferred splicing reaction mixes are described in the Examples below. Methods for the preparation of splicing reaction mixes are provided herein as well as being known in the art. Candidate compounds used in the methods of the invention may be any chemical compound or composition. The description of agonists and antagonists elsewhere herein provide some preferred examples of the kinds of chemical compounds and compositions that may be candidate compounds.

A preferred method includes, for example, *in vitro* splicing using whole cell extracts as reaction mixes prepared from non-pathogens or pathogens, such as the fungi (i) *S. cerevisiae* and, (ii) *Candida albicans* and *Aspergillus fumigatus* respectively. One more preferred embodiment of this assay uses an *S. cerevisiae* actin pre-mRNA in isolated extracts to produce intermediates and products of correct splicing. Another embodiment of this assay uses a pre-mRNA substrate derived from the *C. albicans* beta-tubulin gene that directs accurate splicing in whole cell extracts derived from both *C. albicans* and *S. cerevisiae*. However, any pre-RNA that is capable of mediating a splicing reaction *in vivo*

or *in vitro* may be used in the methods of the invention. Based on the disclosure herein, the skilled artisan could readily choose a splicing reaction mix and splicing pre-RNA to carry out the methods of the invention.

Further, more preferred reaction mixes include, for example, a whole cell extract prepared from *Candida albicans* or *Aspergillus fumigatus* that accurately splices pre-mRNAs derived from unicellular fungi, preferably that of a *S. cerevisiae* actin gene or a *C. albicans* beta-tubulin gene.

The invention also provides a method wherein the modulation is inhibition or activation. Preferred compounds associated with inhibition (antagonism) or activation (agonism) of the splicing reaction will also be anti-pathogen compounds and will be useful in the design of other anti-pathogen compounds. One may demonstrate inhibition or activation of the splicing reaction by comparing the reaction kinetics or reaction products with a control splicing reaction. Inhibition or activation of reaction kinetics or the level of reaction predicts as compared to controls will allow for the identification of the antagonists and agonists respectively.

The invention further provides a method of wherein the splicing reaction comprises at least one component of a pathogen or non-pathogen spliceosome. Spliceosome components required for the splicing reaction have been defined in some model systems and are known in the art. Homologs or orthologs of these components in non-pathogens and pathogens are provided by and useful in the methods of the invention. Preferred spliceosome components include, but are not limited to, snRNPs, particularly those from a pathogen, especially *Candida albicans*. Other components include, for example, pathogen or non-pathogen pre-mRNAs, hn-RNAs, mRNAs

The invention also provides a method of screening for a compound modulating a pathogen or non-pathogen splicing reaction comprising the steps of contacting a splicing reaction mix with a candidate compound and detecting a compound modulating the splicing reaction. A detecting step may be carried out in many ways, some of which will be readily apparent to the skilled artisan based on the invention. Preferred methods of detection include any method that can quantitatively or qualitatively detect any stage of a splicing reaction or the state of any component in a splicing reaction, such as the detection of any form of pre-mRNA (*e.g.* lariats,), calorimetric assays, and gel electrophoresis. Preferred methods of detection provide for the identification RNA structures made during maturation of pathogen or non-pathogen pre-mRNA or hn-RNA to mRNA. Once these RNA structures

are compared to RNA structures from control splicing reactions where no candidate compound was added, one can identify compounds that modulate the maturation of pathogen or non-pathogen pre-mRNA or hn-RNA to mRNA.

A method of inhibiting a splicing reaction is also provided by the invention comprising the step of contacting at least one component of the splicing reaction machinery with a candidate compound of the invention. A component of such machinery is a spliceosome as well as any of the components of the spliceosome, such as snRNPs. Homologs or orthologs of these components or machinery in non-pathogens and pathogens are provided by and useful in the methods of the invention.

Splicing Reaction Polynucleotide and Polypeptides

The invention also relates to splicing polypeptides and polynucleotides and methods of using the same to screen for inhibitors of the splicing reaction. In particular, the invention relates to splicing polypeptides, such as PRP orthologues, and polynucleotides of *Candida albicans*. The invention relates especially to splicing polynucleotide sequences, such as pre-mRNAs, set out in Table 1 as SEQ ID NO: 1 and SEQ ID NO: 2 respectively

---

TABLE 1

Examples of Preferred Splicing Polynucleotide Sequences

(A) *C. albicans* pre-mRNA polynucleotide sequence [SEQ ID NO:1].

5'-CGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCG
GGCTGCAGGAATTCGATAAAGAATTTTTTAATCAAGAAATCAATATCAACATTAT
AGAACTATGAGAGAAATTgtatgttccgtttattcccttccttacaccaaatcattggaatccttatgtgttttgttgttgttgt
ctgaaaatttttggttctgttttacgcgccttttttcgactaaattgattcaaataggattccctaaatgattaatgatttgtgtcaatcaatcaatg
ttttattaagttttatcaaataTACTAACTgcttttgccctttttttttagATTCATTTATCAACTGGTCAATGTGG
TAATCAAATTGTATGTATAAACACTGAAGAAAAAAAAAATTCTATCATTGTTATGT
TGTGATCTTTGATCTTTAGTTGTCGGGTTATCa -3'

(B) *S. cerivisiae* pre-mRNA polynucleotide sequence in this table [SEQ ID NO: 2].

5'-CGAATTGGAGCTcAGCTTTTAGATTTTTCACGCTTACTGCTTTTTTCTTCCCAAG
ATCGAAAATTTACTGAATTAACAATGGattctgGTATGTTCTAGCGCTTGCACCATCC
CATTTAACTGTAAGAAGAATTGCACGGTCCCAATTGCTCGAGAGATTTCTCTTTTA
CCTTTTTTTTACTATTTTTCACTCTCCCATAACCTCCTATATTGACTGATCTGTAATA
ACCACGATATTATTGGAATAAATAGGGGCTTGAAATTTGGAAAAAAAAAAAAAA

CTGAAATATTTTCGTGATAAGTGATAGTGATATTCTTCTTTTATTTGCTACTGTTA
CTAAGTCTCATGtactaacATCGATTGCTTCATTCTTTTTGTTGCTATATTATATGtttagA
GGTTGCTGCTTTGGTTATTGATAACGGTTCTGGTATGTGTAAAGCCGGTTTTGCCG
GTGACGACGCTCCTCGTGCTGTCTTCCCATCTATCGTCGGTAGACCAAGACACCA
AGGTATCATGGTCGGTATGGGTCAAAAAGACTCCTACGTTGGTGATGAAGC
Tg -3'

In one aspect of the invention there are provided polypeptides of a yeast, particularly *Candida albicans* or *Aspergillus fumigatus* referred to herein as "splicing polypeptides," such as PRPs an other protein splicing machinery, as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

In another aspect of the invention there is provided a pre-mRNA of the invention is derived from *Candida albicans* or *Aspergillus fumigatus*, however, it may preferably be obtained from other pathogens or non-pathogens of the same taxonomic genera. A polypeptide of the invention may also be obtained, for example, from pathogen or non-pathogens of the same taxonomic families or orders.

A fragment is a variant polypeptide having an amino acid sequence that is entirely the same as part but not all of any amino acid sequence of any polypeptide of the invention. As with splicing polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides, or of variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, particularly a *Candida albicans* or *Aspergillus fumigatus*, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Also preferred are biologically active fragments which are those fragments that mediate activities of splicing, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those fragments that are antigenic or immunogenic in an animal, especially in a human. Particularly preferred are fragments comprising receptors or domains of enzymes that confer a function essential for viability of *Candida albicans* or *Aspergillus fumigatus* or the ability to initiate, or maintain cause disease in an individual, particularly a human.

It is an object of the invention to provide polynucleotides, such as pre-mRNAs, that serve as a template for a splicing reaction.

In a particularly preferred embodiment of the invention the polynucleotide comprises a region transcribed into a pre-mRNA comprising a sequence set out in Table 1 [SEQ ID NO:1 OR 2], or a variant thereof.

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing splicing polypeptides and polynucleotides, particularly *Candida albicans* or *Aspergillus fumigatus* splicing polypeptides and polynucleotides, including, for example, unprocessed RNAs, ribozyme RNAs, mRNAs, pre-mRNAs, an intron flanked on its 5' and 3' ends by exons, cDNAs, genomic DNAs, B- and Z-DNAs. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful polynucleotides and polypeptides, and variants thereof, and compositions comprising the same.

In another particularly preferred embodiment of the invention there is a splicing polynucleotide, such as a pre-mRNA, from *Candida albicans* or *Aspergillus fumigatus* comprising or consisting of a polynucleotide sequence of Table 1 [SEQ ID NO:1 or 2], or a variant thereof.

In a further aspect, the present invention provides for an isolated polynucleotide, such as a pre-mRNA, comprising or consisting of:

(a) a polynucleotide, such as a pre-mRNA, sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1 or 2 over the entire length of SEQ ID NO:1 or 2 respectively.

Polynucleotides of the present invention, including homologs and orthologs from species other than *Candida albicans* or *Aspergillus fumigatus* that may be obtained by a process which comprises the steps of screening an appropriate library under stringent

hybridization conditions with a labeled or detectable probe consisting of or comprising the sequence of SEQ ID NO: 1 or 2 or a fragment or either; and isolating a polynucleotide that hybridizes thereto.

Polynucleotides, such as pre-mRNAs, of the invention may also contain at least one non-coding sequence, including for example, but not limited to at least one non-coding 5' and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize mRNA, introns, exons, branchpoints, enhancers, splice sites, and polyadenylation signals. The polynucleotide, such as a pre-mRNA, sequence may also comprise additional coding sequence encoding additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA 86:* 821-824 (1989), or an HA peptide tag (Wilson *et al.*, *Cell 37:* 767 (1984), both of which may be useful in purifying polypeptide sequence fused to them. Polynucleotides, such as pre-mRNAs, of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression. Artificial or heterologous sequences that control gene expression, such as sequences and motifs that mediate the splicing reaction, can also be used in the invention.

The invention also includes a polynucleotide, such as a pre-mRNA, consisting of or comprising a polynucleotide of the formula:

$$X\text{-}(R_1)_m\text{-}(R_2)\text{-}(R_3)_n\text{-}Y$$

wherein, at the 5' end of the molecule, X is hydrogen, a metal or a modified nucleotide residue, or together with Y defines a covalent bond, and at the 3' end of the molecule, Y is hydrogen, a metal, or a modified nucleotide residue, or together with X defines the covalent bond, each occurrence of R_1 and R_3 is independently any nucleic acid residue or modified nucleic acid residue, m is an integer between 1 and 3000 or zero , n is an integer between 1 and 3000 or zero, and R_2 is a nucleic acid sequence or modified nucleic acid sequence of the invention, particularly a nucleic acid sequence selected from Table 1 or a modified nucleic acid sequence thereof. In the polynucleotide formula above, R_2 is oriented so that its 5' end nucleic acid residue is at the left, bound to R_1, and its 3' end nucleic acid residue is at the right, bound to R_3. Any stretch of nucleic acid residues denoted by either R_1 and/or R_2, where m and/or n is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer. Where, in a preferred embodiment, X and Y together define a

covalent bond, the polynucleotide of the above formula is a closed, circular polynucleotide, which can be a double-stranded polynucleotide wherein the formula shows a first strand to which the second strand is complementary. In another preferred embodiment m and/or n is an integer between 1 and 1000. Other preferred embodiments of the invention are provided where m is an integer between 1 and 50, 100 or 500, and n is an integer between 1 and 50, 100, or 500.

It is most preferred that a polynucleotide, such as a pre-mRNA, of the invention is derived from *Candida albicans*, however, it may preferably be obtained from other pathogens and non-pathogens of the same taxonomic genus. A polynucleotide, such as a pre-mRNA, of the invention may also be obtained, for example, from pathogens or non-pathogens of the same taxonomic family or order.

The invention further relates to polynucleotides, such as pre-mRNAs, that hybridize to the polynucleotide sequences provided herein. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the polynucleotides described herein. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least 97% identity between the sequences. A specific example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention.

As discussed elsewhere herein regarding polynucleotide assays of the invention, for instance, the polynucleotides of the invention, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones comprising splicing polynucleotides and to isolate cDNA and genomic clones of other genes that have a high identity, particularly high sequence identity, to the splicing polynucleotides. Such probes generally will comprise at least 15 nucleotide residues or base pairs. Preferably, such probes will have at least 30 nucleotide residues or base pairs and may have at least 50 nucleotide residues or base pairs. Particularly preferred probes will have at least 20 nucleotide residues or base pairs and will have lee than 30 nucleotide residues or base pairs.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

For each and every polynucleotide of the invention there is provided a polynucleotide complementary to it. It is preferred that these complementary polynucleotides are fully complementary to each polynucleotide with which they are complementary.

Vectors, Host Cells, Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of splicing polynucleotides or polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems which comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of the polypeptides of the invention, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, and *Streptococcus pneumoniae*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, a basidiomycete, *Candida albicans* and *Aspergillus*; insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

A great variety of expression systems can be used to produce the polypeptides or polynucleotides, such as pre-mRNAs, of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Antibodies

The polypeptides and polynucleotides of the invention or variants thereof, or cells expressing the same can be used as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides respectively.

In certain preferred embodiments of the invention there are provided antibodies against splicing polypeptides or polynucleotides.

Antibodies generated against the polypeptides or polynucleotides of the invention can be obtained by administering the polypeptides and/or polynucleotides of the invention, or

epitope-bearing fragments of either or both, analogues of either or both, or cells expressing either or both, to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature 256:* 495-497 (1975); Kozbor *et al.*, *Immunology Today 4:* 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-splicing or from naive libraries (McCafferty, *et al.*, (1990), Nature 348, 552-554; Marks, *et al.*, (1992) *Biotechnology 10*, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature 352:* 628).

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

Thus, among others, antibodies against splicing-polypeptide or splicing-polynucleotide may be employed to treat infections, particularly fungal infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

A polypeptide or polynucleotide of the invention, such as an antigenically or immunologically equivalent derivative or a fusion protein of the polypeptide is used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin, keyhole limpet haemocyanin or tetanus toxoid. Alternatively, a multiple antigenic polypeptide comprising multiple copies of the polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complimentarily determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

In accordance with an aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization. Among the particularly preferred embodiments of the invention are naturally occurring allelic variants of splicing polynucleotides and polypeptides encoded thereby.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al., Hum Mol Genet* (1992) 1: 363, Manthorpe *et al., Hum. Gene Ther.* (1983) 4: 419), delivery of DNA complexed with specific protein carriers (Wu *et al., J Biol Chem.* (1989) 264: 16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, (1986) 83: 9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al., Science* (1989) 243: 375), particle bombardment (Tang *et al., Nature* (1992) 356:152, Eisenbraun *et al., DNA Cell Biol* (1993) 12: 791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al., PNAS USA* (1984) 81: 5849).

Antagonists and Agonists - Assays and Molecules

Polypeptides and polynucleotides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology 1(2):* Chapter 5 (1991).

Polypeptides and polynucleotides of the present invention are responsible for many biological functions, including many disease states, in particular the Diseases hereinbefore mentioned. It is therefore desirable to devise screening methods to identify compounds which stimulate or which inhibit the function of the polypeptide or polynucleotide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those which stimulate or which inhibit the function of a polypeptide or polynucleotide of the invention, as well as related polypeptides and polynucleotides. In general, agonists or antagonists may be employed for therapeutic and prophylactic purposes for such Diseases as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example,

cells, cell-free preparations, chemical libraries, and natural product mixtures. Such agonists, antagonists or inhibitors so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of splicing polypeptides and polynucleotides; or may be structural or functional mimetics thereof (see Coligan *et al.*, *Current Protocols in Immunology* 1(2):Chapter 5 (1991)).

The screening methods may simply measure the binding of a candidate compound to the polypeptide or polynucleotide, or to cells or membranes bearing the polypeptide or polynucleotide, or a fusion protein of the polypeptide by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide or polynucleotide, using detection systems appropriate to the cells comprising the polypeptide or polynucleotide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptide and/or constitutively expressed polypeptides and polynucleotides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide or polynucleotide, as the case may be. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide or polynucleotide of the present invention, to form a mixture, measuring splicing polypeptide and/or polynucleotide and/or splicing reaction activity in the mixture, and comparing the splicing polypeptide and/or polynucleotide and/or splicing reaction activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and splicing polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists of the polypeptide of the present invention, as well as of phylogenetically and and/or functionally related polypeptides (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies that bind to and/or interact with a polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and/or polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance

the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of splicing polypeptides or polynucleotides, particularly those compounds that are fungistatic and/or fungicidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic splicing reaction mix, a cellular compartment, such as a membrane or cell wall, or a preparation of any thereof, comprising a splicing polynucleotide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate compound that may be a splicing polypeptide and/or polynucleotide and/or splicing reaction agonist or antagonist. The ability of the candidate compound to agonize or antagonize the splicing polypeptide is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously to a component of the splicing machinery, or splicing polynucleotide or polypeptide, *i.e.*, without inducing the effects of splicing reaction are most likely to be good antagonists. Molecules that bind well and, as the case may be, increase the rate of product production from substrate, increase signal transduction, or increase chemical channel activity are agonists. Detection of the rate or level of, as the case may be, production of product from substrate, such as mRNA or otherwise processed RNA from pre-mRNA, may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric, labeled substrate converted into product, a reporter gene that is responsive to changes in splicing polynucleotide or polypeptide or splicing reaction activity, and binding assays known in the art.

Polypeptides of the invention may be used to identify membrane bound or soluble receptors, if any, for such polypeptide, through standard receptor binding techniques known in the art. These techniques include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, ^{125}I), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (*e.g.*, cells, cell membranes, cell supernatants, tissue extracts, bodily materials). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide which compete with the binding of the polypeptide to its receptor(s), if any. Standard methods for conducting such assays are well understood in the art.

In other embodiments of the invention there are provided methods for identifying compounds which bind to or otherwise interact with and inhibit or activate an activity or expression of a polypeptide and/or polynucleotide of the invention comprising: contacting a polypeptide and/or polynucleotide of the invention with a compound to be screened under conditions to permit binding to or other interaction between the compound and the polypeptide and/or polynucleotide to assess the binding to or other interaction with the compound, such binding or interaction preferably being associated with a second component capable of providing a detectable signal in response to the binding or interaction of the polypeptide and/or polynucleotide with the compound; and determining whether the compound binds to or otherwise interacts with and activates or inhibits an activity or expression of the polypeptide and/or polynucleotide by detecting the presence or absence of a signal generated from the binding or interaction of the compound with the polypeptide and/or polynucleotide.

Another example of an assay for splicing reaction agonists is a competitive assay that combines a splicing polynucleotide and a potential agonist with splicing polynucleotide- or polypeptide-binding molecules, recombinant splicing binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. Splicing polynucleotides or polypeptides can be labeled, such as by radioactivity or a colorimetric compound, such that the number of splicing polynucleotides or polypeptides bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing splicing reaction-induced activities, thereby preventing the action or expression of splicing polypeptides and/or polynucleotides by excluding splicing polypeptides and/or polynucleotides from binding.

Small organic molecules of the invention preferably have a molecular weight below 2,000 daltons, more preferably between 300 and 1,000 daltons, and most preferably between 400 and 700 daltons.

Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem. 56:* 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE*

INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of splicing.

Potential antagonists include a small molecule that binds to and occupies a binding site or sequence motif of a splicing polynucleotide or polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented.

Other examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Certain of the splicing polypeptides and polynucleotides of the invention are biomimetics, functional mimetics of the natural splicing polypeptide or polynucleotide. These functional mimetics may be used for, among other things, antagonizing the activity of splicing polypeptide or polynucleotide or as a antigen or immunogen in a manner described elsewhere herein. Functional mimetics of the polypeptides or polynucleotide of the invention include but are not limited to truncated polypeptides or polynucleotide.

In a further aspect, the present invention provides methods of treating abnormal conditions such as, for instance, a disease, related to either an excess of, an under-expression of, an elevated activity of, or a decreased activity of splicing polypeptide and/or polynucleotide and/or a splicing reaction.

If the expression and/or activity of the polypeptide and/or polynucleotide is in excess, several approaches are available. One approach comprises administering to an individual in need thereof an inhibitor compound (antagonist) as herein described, optionally in combination with a pharmaceutically acceptable carrier, in an amount effective to inhibit the function and/or expression of the polypeptide and/or polynucleotide, such as, for example, by blocking the binding of ligands, substrates, receptors, enzymes, etc., or by inhibiting a second signal, and thereby alleviating the abnormal condition. In another approach, soluble forms of the polypeptides still capable of binding the ligand, substrate, enzymes, receptors, etc. in competition with endogenous polypeptide and/or polynucleotide may be administered. Typical examples of such competitors include fragments of the splicing polypeptide and/or polypeptide or components of the splicing reaction.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and

various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

In still another approach, expression of the gene encoding an endogenous splicing polypeptide or polynucleotides can be inhibited using expression blocking techniques. This blocking may be targeted against any step in gene expression, but is preferably targeted against transcription and/or translation and/or the splicing reaction *per se*. An examples of a known technique of this sort involve the use of antisense sequences, either internally generated or separately administered (see, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Alternatively, oligonucleotides which form triple helices with the gene can be supplied (see, for example, Lee *et al., Nucleic Acids Res* (1979) 6:3073; Cooney *et al., Science* (1988) 241:456; Dervan *et al., Science* (1991) 251:1360). These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

Each of the polynucleotide or polypeptide sequences or reaction mixes provided herein may be used in the discovery and development of antifungal compounds. Encoded splicing polypeptides or transcribed pre-RNAs, upon expression, can be used as a target for the screening of antifungal drugs. Additionally, the polynucleotide sequences encoding these splicing sequences or motifs of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

In accordance with yet another aspect of the invention, there are provided splicing agonists and antagonists, preferably fungistatic or fungicidal agonists and antagonists.

The antagonists and agonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

Compositions comprising agonists or antagonists

The invention also relates to compositions comprising the polynucleotides, polypeptides, agonists or antagonists discussed above. Thus, the polypeptides of the present

invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

In therapy or as a prophylactic, the active agent may be administered to a patient as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

Within the indicated dosage range, no adverse toxicological effects are expected with the compounds of the invention which would preclude their administration to suitable patients.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide, polynucleotide, agonist or antagonist, such as the soluble form of a polypeptide, polynucleotide agonist or antagonist of the present invention, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Polypeptides, polynucleotides, agonists and antagonists and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, and the like.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

In-dwelling devices include surgical implants, prosthetic devices and catheters, i.e., devices that are introduced to the body of an individual and remain in position for an extended time. Such devices include, for example, artificial joints, heart valves, pacemakers, vascular grafts, vascular catheters, cerebrospinal fluid shunts, urinary catheters, continuous ambulatory peritoneal dialysis (CAPD) catheters.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of fungi to matrix proteins

exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion. The active agent will preferably be present at a concentration of 1μg/ml to 10mg/ml for bathing of wounds or indwelling devices.

Each reference cited herein is hereby incorporated by reference in its entirety. Moreover, each patent application to which this application claims priority is hereby incorporated by reference in its entirety.

EXAMPLES

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples are illustrative, but do not limit the invention.

Example 1

A set forth above, pre-mRNA splicing provides targets for antifungal drug discovery. Specifically, the Examples provide an *in vitro* pre-mRNA splicing assay from whole cell extracts of *S. cerevisiae* and *Candida albicans* in various screening formats of splicing assays, particularly for HTS.

Whole cell extracts prepared from *Candida albicans* can accurately splice pre-mRNAs derived from the S. cerevisiae actin gene and the *C. albicans* beta-tubulin gene. This assay can be used for HTS of compounds that agonize or antagonize splicing.

Example 2 *in vitro* fungal pre-mRNA splicing system

Suitable substrate pre-mRNAs were examined to determine whether they comprise proper signals for splicing. Lin *et al.* (J. Biol. Chem. **260**:14780-14792, 1985) describe a model yeast pre-mRNA derived from the first intron and flanking exons of the *S. cerevisiae* actin gene (Act1) that is accurately and efficiently spliced in whole cell yeast extracts. The Act1 gene was cloned downstream of a T7 promoter for *in vitro* transcription. Suitable primers were designed, and a 566-bp fragment of the yeast Act1 gene was PCR-amplified directly from a colony of *S. cerevisiae* (strain Abys1) grown on a plate. This fragment was then directionally-cloned into pKSII(+). A suitable pre-mRNA splicing substrate was cloned from *Candida albicans*. A *C. albicans* beta-tubulin gene (Tub1) comprising signals for splicing (see, Smith *et al.*, Gene **63**: 53-63, 1988) was cloned. Moreover, a construct

comprising Tub1 was obtained. A 380-bp band containing the Tub1 gene was removed and cloned downstream of a T7 promoter in pKSII(+) for *in vitro* transcription. A diagram of the two pre-mRNA splicing substrates is shown (exons are represented as boxes, and introns as lines) in Figure 1.

Example 3 Preparation of whole cell extracts

Following the procedures of Lin, *et al.* (J. Biol. Chem. **260**:14780-14792, 1985), whole cell extracts were prepared from both *S. cerevisiae* and *C. albicans*. Briefly, 1-2 L of cells were pelleted and resuspended in sorbitol buffer. Treatment of the resuspended cells with zymolase resulted in the formation of spherolpasts. After lysis of the spheroplasts by Dounce homogenation, the lysate was centrifuged several times to remove insoluble material and then dialyzed in storage buffer.

Example 4 Testing of whole cell extracts for splicing activity

Testing *S. cerevisiae* and *C. albicans* whole cell extracts for *in vitro* splicing activity comprised using radiolabeled Act1 pre-mRNA transcribed *in vitro* and column-purified. Labeled pre-mRNA was incubated in either the *S. cerevisiae* or *C. albicans* whole cell extracts under standard splicing conditions (see, for example, Lin, *et al.*, J. Biol. Chem. **260**:14780-14792, 1985). After incubation, the reactions were deproteinated with Proteinase K, phenol extracted, and the RNAs were ethanol precipitated. Electrophoresis on an 8% acrylamide-7M urea gel, followed by autoradiography, allowed direct visualization of intermediates and products of the *in vitro* splicing reaction. Preliminary results indicate that although there is noticeable degradation of the pre-mRNA in the cell extract, the expected intermediates and products of the splicing reaction are evident (based on their sizes), indicating that splicing is indeed occurring. It was demonstrated that *S. cerevisiae* Act1 pre-mRNA is believed to be spliced accurately in *C. albicans* extract.

To examine if the *C. albicans* beta-Tub pre-mRNA would be spliced accurately *in vitro*, 15 min. and 45 min. splicing reactions were performed and analyzed via gel electrophoresis. This result is shown in Figure 2.

Example 5 Splicing timecourse analysis

A timecourse of splicing on the beta-Tub pre-mRNA in whole cell extracts of *C. albicans* clearly shows the appearance of lariat intermediates and a band corresponding to fully-spliced mRNA (note that free lariat and lariat-exon 2 intermediates do not migrate true

to their predicted sizes because these circular molecules are known to migrate anomalously in acrylamide gels). This provide evidence of a reproducible *in vitro* pre-mRNA splicing in *C. albicans*. Importantly, the establishment of an *in vitro* assay to examine splicing of fungal pre-mRNAs provides a basis for the compound screening assays, HTS and otherwise, provided herein.

Example 6 Splicing complex assembly

From the splicing gels, it was determined that both Act1 and beta-Tub pre-mRNAs are spliced accurately in cell extracts from both *S. cerevisiae* and *C. albicans*. To look at the formation of splicing complexes (spliceosomes), a native gel system was adapted from Pikielny and Rosbash (Cell **45**: 869-877, 1986). A timecourse of *in vitro* splicing was performed in *C. albicans* whole cell extracts, and the reactions were stopped and loaded directly onto a native gel (3% acrylamide, 0.5% agarose, 5% glycerol in 0.5x TBE). This analysis resulted in the data in Figure 3. Complexes are identified according to the standard nomenclature for this gel (Pikielny and Rosbash, Cell **45**: 869-877, 1986). The timecourse shows the appearance of the pre-spliceosome and then the movement of the pre-mRNA into the mature spliceosome. Surprisingly, *C. albicans* beta-Tub pre-mRNA appears to be a better substrate for splicing complex assembly then does *S. cerevisiae* Act1 pre-mRNA (as determined by a comparison of the kinetics of appearance of the complexes). This is consistent with analysis of the splicing of the different pre-mRNAs. This analysis of splicing complex assembly provides a compound screening embodiment useful as a secondary assay for, among other things, SAR to examine the mechanistic effects of different agonists and antagonists.

Using methods provided herein, splicing is compared in whole cell extracts prepared by various methods to optimize splicing reactions and to make alternative assay formats.

What is claimed is:

1. A method of screening for a compound modulating a pathogen splicing reaction comprising the steps of contacting a splicing reaction mix with a candidate compound and identifying a compound modulating the splicing reaction.

2. The method of claim 1 wherein the modulating is inhibition.

3. The method of claim 1 wherein the modulating is activation.

4. The method of claim 1 wherein the pathogen is selected from the group consisting of a fungus, protozoan, nematode and arthropod.

5. The method of claim 1 wherein the splicing reaction comprises at least one component of a pathogen splicesome.

6. A method of screening for a compound modulating a pathogen splicing reaction comprising the steps of contacting a splicing reaction mix with a candidate compound and detecting a compound modulating the splicing reaction.

7. A compound identified using the method of claim 1 or selected using the method of claim 7.

8. A method of inhibiting a splicing reaction comprising the step of contacting the reaction machinery with a compound of claim 7.

9. The method of inhibiting the splicing reaction of claim 8 wherein the reaction is in a pathogen..

10. A method of treating an individual infected with a pathogen comprising the step of contacting the individual with the a compound of claim 7.

11. A method of treating an individual infected with a pathogen comprising the step of contacting the individual with the a compound that modulates the activity of a splicing reaction in a pathogen.

12. An antagonist or agonist of a pathogen splicing reaction.

13. The method of claim 4 wherein the fungus is selected form the group consisting of S. cerevisiae and C. albicans.

14. A method of screening for a compound modulating a the maturation of pathogen pre-mRNA or hn-RNA to mRNA comprising the steps of contacting a composition comprising a pre-mRNA or hn-RNA with a candidate compound and identifying a compound modulating the maturation of pathogen pre-mRNA or hn-RNA to mRNA.

15. A compound identified using the method of claim 14.

16. A method of modulating the maturation of pathogen pre-mRNA or hn-RNA to mRNA with a compound of claim 14.

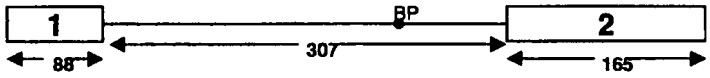
17. A method of treating an individual infected with a pathogen comprising the step of contacting the individual with the a compound identified using the method of claim 14.

18. A method of treating an individual infected with a pathogen comprising the step of contacting the individual with a compound that modulates the maturation of pre-mRNA or hn-RNA to mRNA.

19. A method of treating an individual infected with a pathogen comprising the step of contacting the individual with the compound of claim 14.

FIGURE 1

Act1 (*S. cerevisiae*)



Tub1 (*C. albicans*)

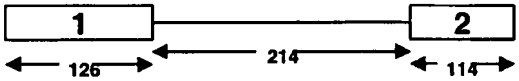


FIGURE 2

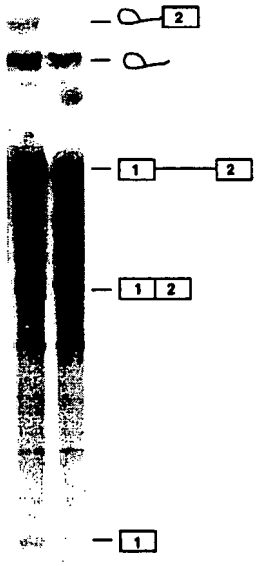
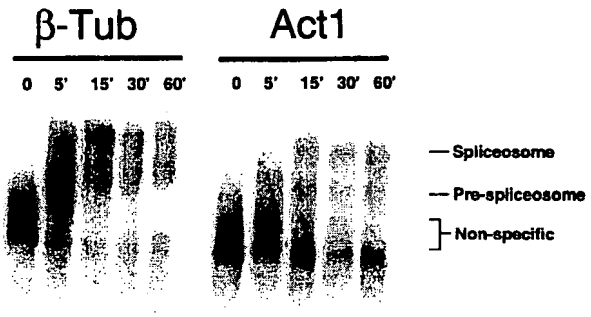


FIGURE 3


β-Tub
Act1
0 5' 15' 30' 60'
0 5' 15' 30' 60'
Spliceosome
Pre-spliceosome
Non-specific

SEQUENCE LISTING

<110> SMITHKLINE BEECHAM CORPORATION
SMITHKLINE BEECHAM plc

<120> METHOD FOR IDENTIFYING COMPOUNDS THAT MODULATE EUKARYOTIC SPLICING

<130> GM50052

<140> TO BE ASSIGNED
<141> 2000-05-04

<150> US 60/133,001
<151> 1999-05-07

<160> 2

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 454
<212> DNA
<213> Candida albicans

<400> 1

cgaattggag ctccaccgcg gtggcggccg ctctagaact agtggatccc ccgggctgca 60
ggaattcgat aaagaatttt ttaatcaaga aatcaatatc aacattatag aactatgaga 120
gaaattgtat gttccgttta ttcccttcct tacaccaaat cattggaatc cttatgtgtt 180
ttgttgttgt tgtctgaaaa tttttggttc tgttttacgc gccttttttcg actaaattga 240
ttcaaatagg attccctaaa tgattaatga tttgtgtcaa tcaatcaatg ttttattaag 300
ttttatcaaa tatactaact gcttttgccc tttttttttag attcatttat caactggtca 360
atgtggtaat caaattgtat gtataaacac tgaagaaaaa aaaattctat cattgttatg 420
ttgtgatctt tgatctttag ttgtcgggtt atca 454

<210> 2

<211> 562
<212> DNA
<213> Saccharomyces cerevisiae

<400> 2

cgaattggag ctcagctttt agatttttca cgcttactgc ttttttcttc ccaagatcga 60
aaatttactg aattaacaat ggattctggt atgttctagc gcttgcacca tcccatttaa 120
ctgtaagaag aattgcacgg tcccaattgc tcgagagatt tctcttttac ctttttttac 180
tatttttcac tctcccataa cctcctatat tgactgatct gtaataacca cgatattatt 240
ggaataaata ggggcttgaa atttggaaaa aaaaaaaaaa ctgaaatatt ttcgtgataa 300
gtgatagtga tattcttctt ttatttgcta ctgttactaa gtctcatgta ctaacatcga 360
ttgcttcatt ctttttgttg ctatattata tgtttagagg ttgctgcttt ggttattgat 420
aacggttctg gtatgtgtaa agccggtttt gccggtgacg acgctcctcg tgctgtcttc 480
ccatctatcg tcggtagacc aagacaccaa ggtatcatgg tcggtatggg tcaaaaagac 540
tcctacgttg gtgatgaagc tg 562

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/12249

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : A01N 61/00; A61K 31/00; C12P 19/34; C12Q 1/68
US CL : 435/6, 91.31; 514/1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6, 32, 375, 91.31; 514/1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 5,900,406 A (VON AHSEN et al) 04 May 1999. See column 2, line 16, to column 8, line 20; especially column 3, line 6, to column 4, line 29. Also not claim 1-3 and 8-10.	1-12, 14-19 ---------- 13
X --- Y	US 5,849,484 A (LEIBOWITZ et al.) 15 December 1998. See column 5, line 45, to column 7, line 6; column 9, line 3, to column 11, line 55.	1-12, 14-19 ---------- 13
X	SERAPHIN et al. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. The EMBO Journal. 1988, Vol. 7, No. 8, pages 2533-2538, especially abstract and page 2537, column 1, first and second full paragraphs.	1-10, 12-16
Y	Database Medline on STN, US National Library of Medicine, (Bethesda MD, USA), No. 91031510, SILVER et al. 'Screening of natural products for antimicrobial agents.' abstract, European Journal of Clinical Microbiology and Infectious Diseases (July 1990)Vol. 9, No. 7, pages 455-461.	1-6, 14

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 July 2000 (31.07.2000)	24 AUG 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer THOMAS G. LARSON, Ph.D. Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/12249

Continuation of B. FIELDS SEARCHED Item 3: USPAT, DWPI, EPOABS, JPABS (all using WEST), BIOSIS, CAPLUS, MEDLINE (all using STN). Key words: splicing, splicesome, pre-RNA, hn-RNA, mRNA, maturation, modulate, inhibit, stimulate, screen, pathogen, yeast, fungus, protozoan, nematode, arthropod, snrnp,